MATLAB EXPO 2018

Riutilizzo e prototipazione di codice
Design of Voice Interfaces for IoT Devices

Francesca Perino
- Innovate
- Reuse
- Prototype
What Device Is This?
What Are Microphone Arrays?
Why Microphone Arrays?
good afternoon everybody
good afternoon everybody
do you speak Matlab
How Can I…

1. Design a microphone array system?

2. Validate my voice interface can work in real-life scenarios?

3. Understand what else can help me improve my performance?
How Can I…

1. Design a microphone array system?

2. Validate my voice interface can work in real-life scenarios?

3. Understand what else can help me improve my performance?
Design a Microphone Array System
...starting from a given array hardware

- BlueCoin from ST Microelectronics
Azimuth Cut (elevation angle = 0.0°)

Directivity (dBi), Broadside at 0.00 degrees
Time Domain Simulation of an Array System
% Microphone pair
ula = phased.ULA(2, dx, 'Taper', [1 -1]);

% Acquisition of propagating signal
collector = phased.WidebandCollector('Sensor', ula,...
    'PropagationSpeed', c,'SampleRate', fs,'ModulatedInput',false,...
    'NumSubbands',4096);
incidentAngle = [toTune.sourceTheta;0];

% Beamformer
beamform = phased.TimeDelayBeamformer('SensorArray', ula,...
    'SampleRate', fs,'PropagationSpeed',c,...
    'DirectionSource', 'Input port', 'WeightsOutputPort',false);
nullSteeringAngle = [toTune.steeringTheta;0];

% Spectrum analysis
spect = dsp.SpectrumAnalyzer('SampleRate', fs,'PlotAsTwoSidedSpectrum',false,...
    'YLimits', [-50,15], 'ShowLegend', true,...
    'SpectralAverages', 128, 'FrequencyScale', 'Log');
show(spect)
adjustPositions(viewer, spect)
The Spectrum Analyzer requires 3073 samples to update the display. It will update once the required number of samples have been input.
How Can I ... Validate my voice interface can work in real-life scenarios?
Constrained Simulations vs. Real Life
release(readaudio)

%% Stream and visualize live

% Scope for time-domain visualization
tscope = dsp.TimeScope('SampleRate',fs,'TimeSpan',0.500,...
    'YLimits',[-0.5,0.5],'TimeSpanOverrunAction','Scroll');

% Set shorter block duration
readaudio.SamplesPerFrame = 1024;

tic
while toc < 30
    % Acquire live
    in = readaudio();
    % Visualize in real-time
tscope(in);
end

%%
release(readaudio)

>> readaudio.Device = 'Microphone (STM32 AUDIO Streaming in FS Mode)'
```matlab
%codegen

% properties

% "Looking" direction
Theta = 0
% Bandpass filter gain (dB) - 0=all bandpassed / -inf=allpass
BandpassGain = 0
% Beam type
Beamtype = BeamformingType.Cardiod
% Output Gain (dB)
OutputGain = 0

end

properties (Access = private)

th = 0

beamformingType = 3
```
```matlab
classdef LevelAndBeamform < audioPlugin

    % LevelAndBeamform with properties:
    %
    %   Theta: 0
    %   BandPassGain: 0
    %   BeamType: Cardioid
    %   OutputGain: 12

    >> audioTestBench(voiceProc)
```
How can I…

2. Validate my voice interface can work in real-life scenarios?

3. Understand what else can help me improve my performance?
Testing an external audio plugin within a MATLAB live system model.

The Spectrum Analyzer requires 3073 samples to update the display. It will update once the required number of samples have been input.
Plugin hosting

```matlab
>> noiseRemover = loadAudioPlugin('ERA-N.vst')
noiseRemover =

VST plugin 'ERA-N' 2 in, 2 out

Processing: 40 %
Gain: 0 dB
Tilt: 'NoTilt'
Bypass: 0

>> noiseRemover.Processing = 60;
>> noiseRemover.Gain = 3;
>> y = process(noiseRemover, x)
```

https://accusonus.com/products/era-n
How To Measure Performance?

Output audio "sounds good"

"91.5% of spoken sentences correctly converted"
Test performance with speech-to-text services

```matlab
>> [samples, fs] = audioread('helloaudioPD.wav');
>> soundsc(samples, fs)
>> speechObject = speechClient('Google','languageCode','en-US');
>> outInfo = speech2text(speechObject, samples, fs);

>> outInfo.TRANSCRIPT =
ans =
    'hello audio product Developers'

>> outInfo.CONFIDENCE =
ans =
    0.9385
```

https://www.mathworks.com/matlabcentral/fileexchange/65266-speech2text

MATLAB EXPO 2018
Building a small speech dataset quickly

Apps enable interactivity and automation.
Building a small speech dataset quickly
How to accelerate speech content labelling?

http://www.cs.cmu.edu/afs/cs.cmu.edu/project/lddata/oldfiles/Recorder.app/utterances/Type1/harvsents.txt
Building a small speech dataset quickly
Example: an App with automated content labelling
Building a small speech dataset quickly

Example: an App with automated content labelling*

*See also Dataset Recorder App prototype in example "Record Audio Datasets" (From R2018a in Audio System Toolbox)
%% Cycle through recordings

% Init arrays to store results
numFiles = length(ds.Files);
k = 0;

conf = zeros(1,numFiles);
trueStrings = repmat('"',1,1,numFiles);
guessedStrings = repmat('"',1,1,numFiles);

while hasdata(ds)
    % Read audio file content
    entry = read(ds);
    x = entry.samples;
    fprintf('Sending "%s"...
', entry.labels.SpeechContent)

    % Send ORIGINAL audio to cloud service and get transcription
    [outString, outInfo] = speech2text(x(:,1), entry.fs);

    % Store results
    k = k + 1;
    conf(k) = outInfo.confidence;
    trueStrings(k) = entry.labels.SpeechContent;
    guessedStrings(k) = outString;

end

tResults = table(trueStrings', guessedStrings', num2str(conf)',
    'VariableNames', {'True', 'Guessed', 'Confidence'})
It's easy to tell the depth of a well.

These days a chicken leg is a rare dish.

Rice is often served in round bowls.

It's easy to tell the depth of a well.

These days a chicken leg is a rare dish.

In a large size in stockings, it's hard to sell.

A large size in stockings, it's hard to sell.

A large size in stockings is hard to sell.

A large size in stockings, it's hard to sell.
How Can I…

1. Design a microphone array system?

2. Validate my voice interface can work in real-life scenarios?

3. Understand what else can help me improve my performance?
Summary

- Innovate
- Reuse
- Prototype